Report on Python Group Project——Stock Picker

Fangjun Liu (fl2413)   Yunfan Zhang (yz2866)   Zhuangfei Yang (zy2233)   Zimeng Ji (zj2194)

1. Introduction

In this report, we will mainly focus on the explanation of our work. For the sake of the report, detailed graphs and outputs are included in the submitted files.

2. Dataset

We included the daily adjusted closing prices starting from 11/18/2010 to 7/1/2015 of 144 stocks (top 16 stocks in 9 industries) as well as SP 500 grabbed from Yahoo Finance in our dataset.

3. Principal work

3.1 Basic Analysis

We did some basic analysis including the calculation of the mean and variance of return rates of each stock in different industries and graphed the results.

Then, we calculated the beta for each stock. We found that the beta values of stocks of periodic industries (beta>1) were always higher than the beta values of stocks of non-periodic industries (beta<0.7). This finding actually corresponds to the proven financial rule.

3.1 Stock Picking Strategy I

For the first strategy, our program will create a portfolio for the clients each they run our script. The basic concept is to use two different formulas to grade stocks during different market scenarios, pick up the top three stocks with highest scores in each sector to create the portfolio. We used data from our data set to build the portfolio and used the recent 5 months' data to test the performance of this strategy. The test results shown in our program proved that the two portfolios can both beat the market regardless of whether the market is going-up or going-down.

3.2 Stock Picking Strategy II

The second strategy we designed will automatically generate the top ten stocks of SP 500 every week and invest equal amount of money in each of them next week. The total return from 2011 until recently is 103.79%. The chart indicating the performance of the portfolio created by strategy II in comparison to S&P 500 index which you can see in our program distinctly proved the effectiveness of this strategy.

3.3 Study on the effect of the number of news on market forecasting

We proceeded to research whether the number of news can predict the market to some extent. The quantities of news are automatically calculated through accessing the website of NewYorkTimes. We used the Selenium to do achieve this since we could not get the complete data set via BeautifulSoup.

After testing and scanning many financial news, we selected two sets of common words for positive news and negative news respectively as searching strings. That is, we used 'rebound', 'stimulate' and 'boost', etc. to search positive news, and we used 'breakdown', 'slowdown' and 'suffer', etc. to search negative news. Search strings can be adjusted in further research.

After carefully observing the correlation between the number of news and future returns of SP 500 under different forecast periods. We found that the one-month-later returns of SP 500 has relatively higher correlation with the number of news. Then we plotted the number of negative and positive news against the one-month-later returns of SP 500 (correspondingly enlarged to plot in one graph) in 2011. We further found that the number of positive news matched better than the number of negative news with returns. Through plotting these data during 2012-2015, we found these two variables matched well at significant numbers of peaks and bottom. Therefore, we could use the number of positive news one month before as an alternative factor to forecast today's market and possibly utilize this pattern to create strategies to earn larger profit.